

# Extended Non-Binary Low-Density Parity-Check Codes over Erasure Channels

Lam PHAM SY  
EUTELSAT SA  
F-75015, Paris, France  
lphamsy@eutelsat.fr

Valentin SAVIN  
CEA-LETI, MINATEC campus  
F-38054, Grenoble, France  
valentin.savin@cea.fr

David DECLERCQ  
ETIS, ENSEA / Univ. Cergy-Pontoise / CNRS UMR-8051  
F-95000, Cergy-Pontoise, France  
declercq@ensea.fr

**Abstract**—Based on the extended binary image of non-binary LDPC codes, we propose a method for generating extra redundant bits, such as to decrease the coding rate of a mother code. The proposed method allows for using the same decoder, regardless of how many extra redundant bits have been produced, which considerably increases the flexibility of the system without significantly increasing its complexity. Extended codes are also optimized for the binary erasure channel, by using density evolution methods. Nevertheless, the results presented in this paper can easily be extrapolated to more general channel models.

**Index Terms**—Non-binary LDPC codes, extended binary image, incremental redundancy, very small coding rates.

## I. INTRODUCTION

Adaptive coding techniques are frequently employed, especially in wireless communications, in order to dynamically adjust the coding rate to changing channel conditions. An example of adaptive coding technique consists in puncturing a mother code. When the channel conditions are good more bits are punctured and the coding rate is increased. In poor channel conditions all redundant bits are transmitted and the coding rate drops. However, in harsh conditions, the receiver might not be able to successfully decode the received signal, even if all the redundant bits have been transmitted. In such a case, the coded block can be retransmitted until the sent information is successfully decoded. This is equivalent to additional repetition coding, which further lowers the coding rate below the mother coding rate.

However, the use of retransmission techniques might not be suitable nor possible in some situations, such as multi-cast/broadcast transmissions, or whenever the return link is strictly limited or not available (such situations are generally encountered in satellite communications). The main alternative in this case is the use of *erasure codes* that operate at the transport or the application layer of the communication system: source data packets are extended with redundant (also referred to as *repair*) packets that are used to recover the lost data at the receiver. Physical (PHY) and upper layer (UL) codes are not mutually exclusive, but they are complementary to each other. Adaptive coding schemes are also required at the upper layer, in order to dynamically adjust to variable loss rates. Besides, codes with very small rates or even *rateless* [1], [2] are sometimes used at the application layer for fountain-like content distribution applications.

This work was supported by the French National Research Agency (ANR), grant No 2009 VERS 019 04 – ARSSO project.

In this paper we propose a coding technique that allows to produce extra redundant bits, such as to decrease the coding rate below the mother coding rate. Extra redundant bits can be produced in an incremental way, yielding very small coding rates, or can be optimized for a given target rate below the mother coding rate. As for puncturing, the proposed technique allows for using the same decoder, regardless of how many extra redundant bits have been produced, which considerably increases the flexibility of the system, without increasing its complexity.

The proposed coding scheme is based on non-binary low density parity check (NB-LDPC) codes [3] or, more precisely, on their *extended binary image* [4]. If  $q = 2^p$  denotes the size of the non-binary alphabet, each non-binary symbol corresponds to a  $p$ -tuple of bits, referred to as its binary image. Extra redundant bits, called *extended bits*, are generated as the XOR of some bits from the binary image of the same non-binary coded symbol. If a certain number of extended bits are transmitted over the channel, we obtain an *extended code*, the coding rate of which is referred to as *extended (coding) rate*. In the extreme case when all the extended bits are transmitted, the mother code is turned into a *very small rate* code, and can be used for fountain-like content distribution applications [4]. A similar approach to fountain codes, by using multiplicatively repeated NB-LDPC codes, has been proposed in [5]. If some extended rate is targeted, we show that the extended code can be optimized by using density evolution methods.

The paper is organized as follows. Section II gives the basic definitions and the notation related to NB-LDPC codes. In Section III, we introduce the extended NB-LDPC codes and discuss their erasure decoding. The analysis and optimization of extended NB-LDPC codes are addressed in Section IV. Section V focuses on the code design and presents simulation results, and Section VI concludes the paper.

## II. NON-BINARY LDPC CODES

We consider NB-LDPC codes defined over  $\mathbb{F}_q$  [6], the finite field with  $q$  elements, where  $q = 2^p$  is a power of 2 (this condition is only assumed for practical reasons). We fix once for all an isomorphism of vector spaces:

$$\mathbb{F}_q \xrightarrow{\sim} \mathbb{F}_2^p$$

Elements of  $\mathbb{F}_q$  will also be referred to as *symbols*, and we say that  $\underline{x} = (x_0, \dots, x_{p-1}) \in \mathbb{F}_2^p$  is the *binary image* of the symbol  $X \in \mathbb{F}_q$ , if they correspond to each other

by the above isomorphism. A non-binary LDPC code over  $\mathbb{F}_q$  is defined as the kernel of a sparse parity-check matrix  $H \in \mathbf{M}_{M,N}(\mathbb{F}_q)$ . Alternatively, it can be represented by a bipartite (Tanner) graph [7] containing symbol-nodes and constraint-nodes associated respectively with the  $N$  columns and  $M$  rows of  $H$ . A symbol-node and a constraint-node are connected by an edge if and only if the corresponding entry of  $H$  is non-zero; in this case, the edge is assumed to be *labeled* by the non-zero entry. As usually [8], we denote by  $\lambda$  and  $\rho$  the left (symbol) and right (constraint) edge-perspective degree distribution polynomials. Hence,  $\lambda(x) = \sum_d \lambda_d x^{d-1}$  and  $\rho(x) = \sum_d \rho_d x^{d-1}$ , where  $\lambda_d$  and  $\rho_d$  represent the fraction of edges connected respectively to symbol and constraint nodes of degree- $d$ . The design coding rate is defined as  $r = 1 - \frac{\int_0^1 \rho(x) dx}{\int_0^1 \lambda(x) dx}$ , and it is equal to the coding rate if and only if the parity-check matrix is full-rank.

### III. EXTENDED NON-BINARY LDPC CODES

#### A. Extended code description

For any integer  $1 \leq k \leq q-1$ , let  $[k] = (k_0, \dots, k_{p-1})^T$  denote the column vector corresponding to the binary decomposition of  $k$ ; i.e.  $k = \sum_{i=0}^{p-1} k_i 2^i$ , with  $k_i \in \{0, 1\}$ . Let  $X \in \mathbb{F}_q$  be a non-binary symbol, and  $\underline{x} = (x_0, \dots, x_{p-1})$  be its binary image. The  $k$ -th *extended bit* of  $X$  is by definition:

$$\alpha_k = \underline{x} \times [k] = \sum_{i=0}^{p-1} k_i x_i$$

The vector  $\underline{\alpha} = (\alpha_1, \dots, \alpha_{q-1})$  is called *extended binary image* of  $X$ . Note that  $\alpha_{2^i} = x_i$ , for any  $0 \leq i \leq p-1$ . An extended bit  $\alpha_k$  is said to be *nontrivial* if  $k$  is not a power of 2 (hence,  $\alpha_k$  is a linear combination of at least two bits from the binary image  $\underline{x}$  of  $X$ ).

Now, consider a NB-LDPC code defined over  $\mathbb{F}_q$ , with coding rate  $r = \frac{K}{N}$ . Let  $(X_1, \dots, X_N) \in \mathbb{F}_q^N$  be a non-binary codeword,  $(\underline{x}_1, \dots, \underline{x}_N) \in \mathbb{F}_2^{Np}$  be its binary image, and  $(\underline{\alpha}_1, \dots, \underline{\alpha}_N) \in \mathbb{F}_2^{N(q-1)}$  be its extended binary image. By transmitting the extended binary image over the channel, we obtain a code with rate  $\frac{Kp}{N(q-1)} = r \frac{p}{q-1}$ , which can be advantageously used for application requiring very small coding rates [4].

We define an *extension of the NB-LDPC code*, as a family of matrices  $\{A_1, \dots, A_N\}$ , where each  $A_n \in \mathbf{M}_{p,t_n}(\mathbb{F}_2)$  is a binary matrix with  $p$  rows and  $t_n$  columns, with  $t_n \geq 0$ . Let  $\underline{a}_n = \underline{x}_n \times A_n \in \mathbb{F}_2^{t_n}$ ; hence,  $\underline{a}_n$  is constituted of  $t_n$  extended bits of  $X_n$  (possibly with repetitions, if  $A_n$  contains two or more identical columns). The binary vector  $(\underline{a}_1, \dots, \underline{a}_N)$  is called *extended codeword*, and the *extended coding rate* is given by  $r_e = \frac{Kp}{T}$ , where  $T = \sum_{n=1}^N t_n$ .

Note that the above definition is very broad, and it can yield extended rates below as well as extended rates above the mother coding rate. In particular, it includes punctured codes: if  $A_n = \text{col}([2^1], \dots, [2^{p-1}])$ , then  $\underline{a}_n = (x_{n,1}, \dots, x_{n,p-1})$ , which is the same as puncturing the first bit,  $x_{n,0}$ , from the binary image of  $X_n$ . Moreover, taking some  $t_n = 0$  is

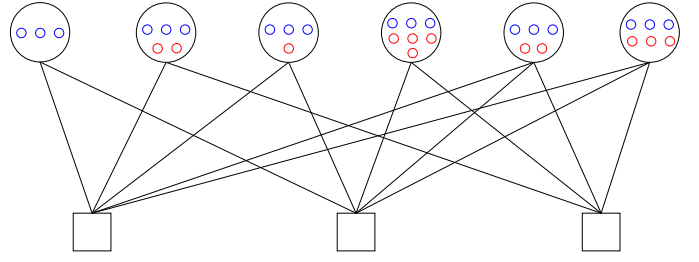


Figure 1. Extension of a NB-LDPC code. Blue circles represent bits of the binary image ( $p = 3$ ), while red circles represent (nontrivial) extended bits.

equivalent to puncturing the whole symbol  $X_n$ . The optimization of puncturing distributions for NB-LDPC codes has been addressed in [9]. In this paper, we restrict ourselves to the case when matrices  $A_n$  are of the form  $A_n = [I_p \mid B_n]$ , where  $I_p$  is the  $p \times p$  identity matrix, meaning that each  $\underline{a}_n$  contains the binary image  $\underline{x}_n$  (the use of “extension” complies with its literal meaning). We will further assume that any two columns of a matrix  $A_n$  are different. It follows that  $p \leq t_n \leq q-1$ , and each  $\underline{a}_n$  is constituted of the binary image  $\underline{x}_n$  and  $t_n - p$  pairwise different (nontrivial) extended bits. In this case, we shall say that *the symbol  $X_n$  is extended by  $k = t_n - p$  bits*.

For instance, let  $p = 3$  and consider that the following matrix  $A$  is used to extend by 2 bits some coded symbol  $X$ :

$$\underline{a} = \underline{x} \times A, \text{ where } A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Then  $\underline{a} = (\alpha_1, \alpha_2, \alpha_4, \alpha_5, \alpha_6) = (x_0, x_1, x_2, x_0 \wedge x_2, x_1 \wedge x_2)$ , where  $\wedge$  denotes the bit-xor operator.

In order to determine the coding rate of the extended code, we denote by  $f_{d,k}$  the *fraction of degree- $d$  symbols with  $k$  nontrivial extended bits*,  $0 \leq k < q-p$ ; thus  $\sum_{k=0}^{q-p-1} f_{d,k} = 1$ . The average number of nontrivial extended bits per coded symbol is given by  $f = \sum_{d=1}^{d_s} \Lambda_d \sum_{k=0}^{q-p-1} k f_{d,k}$ , where  $d_s$  is the maximum symbol node degree, and  $\Lambda_d = \frac{\lambda_d}{d \int_0^1 \lambda(x) dx}$  is the fraction of degree- $d$  symbol nodes. It follows that the extended coding rate is given by  $r_e = r \frac{p}{p+f}$ . Thus, we can achieve an arbitrary extended rate within the interval  $\left[r \frac{p}{q-1}, r\right]$  by varying the parameter  $f$ .

Figure 1 illustrates an extended code defined over  $\mathbb{F}_8$ , with  $f_{2,0} = f_{2,1} = f_{2,2} = f_{2,4} = 1/4$ , and  $f_{3,2} = f_{3,3} = 1/2$ , which correspond to  $f = 2$ . The mother coding rate is  $r = 0.5$  and the extended coding rate is  $r_e = 0.3$ .

#### B. Iterative erasure decoding

We consider that the extended codeword  $(\underline{a}_1, \dots, \underline{a}_N)$  is transmitted over a binary erasure channel (BEC). At the receiver part, the received bits (both from the binary image and extended bits) are used to reconstruct the corresponding non-binary symbols. Precisely, for each received bit we know its position within the extended binary image of the corresponding symbol. Hence, for each symbol node we can determine a set of *eligible symbols* that is constituted of symbols whose extended binary images match the received bits. These sets

are then iteratively updated, according to the linear constraints between symbol-nodes [10]. Alternatively (and equivalently), the extended code can be decoded by using the linear-time erasure decoding proposed in [4].

The *asymptotic threshold* of an ensemble of codes is defined as the maximum erasure probability  $p_{th}$  that allows transmission with an arbitrary small error probability when the code length tends to infinity [8]. Given an ensemble of codes, its threshold value can be efficiently computed by tracking the fraction of erased messages passed during the belief propagation decoding; this method is referred to as *density evolution*. In this paper, the density evolution is approximated based on the Monte-Carlo simulation of an infinite code, similar to the method presented in [9]. This method has two main advantages: it can easily incorporate the extending distribution  $\{f_{d,k}\}_{d,k}$ , and it can be extrapolated to more general channel models.

#### IV. ANALYSIS AND OPTIMIZATION

The goal of this section is to answer the following questions.

First of all, assume that we have given a symbol node that has to be extended by  $k$  bits. How should they be chosen among the  $q - p - 1$  (nontrivial) extended bits?

Secondly, given an extended coding rate  $r_e$ , how should be extended bits distributed over the symbol-nodes? Put differently, which is the optimal extending distribution  $\{f_{d,k}\}$ ?

##### A. Extended bits selection strategy

We assume that we have given a symbol-node that has to be extended by  $k$  bits. A choice of the  $k$  bits among the  $q - p - 1$  extended bits corresponds to an extending matrix  $A = [I_p \mid B]$  of size  $p \times (p + k)$ , with pairwise distinct columns. For each such a matrix, assume that the extended symbol  $\underline{a} = \underline{x} \times A$  is transmitted over the BEC, and let  $E(A)$  be the expected number of eligible symbols at the receiver. Recall that an eligible symbol is a symbol whose extended binary image match the received bits. If all transmitted bits have been erased, any symbol is eligible. Conversely, if the received bits completely determine the non-binary symbol,

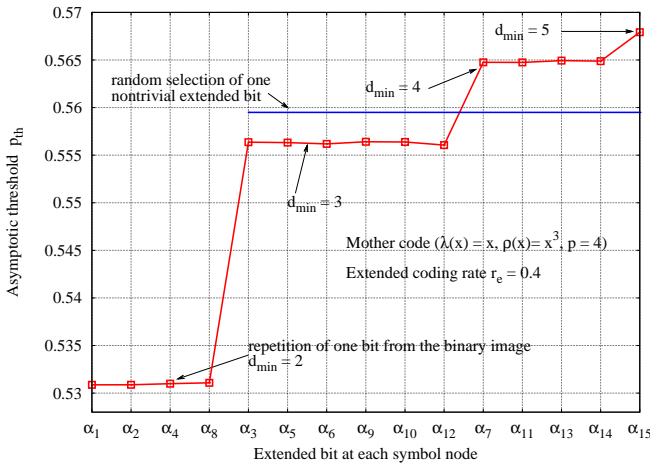


Figure 2. 1-bit extension for regular NB-LDPC codes over  $\mathbb{F}_{16}$

then there is only one eligible symbol. More generally, let  $\underline{a}_{rec}$  denote the sequence of received bits, and  $A_{rec}$  denote the submatrix of  $A$  determined by the columns that correspond to the received positions of  $\underline{a}$ . Then the eligible symbols are the solutions of the linear system  $\underline{x} \times A_{rec} = \underline{a}_{rec}$ , and their number is equal to  $2^{p - \text{rank}(A_{rec})}$ . Now, if  $\epsilon$  denotes the erasure probability of the BEC, it can be easily verified that:

$$E(A) = \sum_{i=0}^{p+k} (1 - \epsilon)^i \epsilon^{p+k-i} \left( \sum_{A_i \subseteq A} 2^{p - \text{rank}(A_i)} \right),$$

where the second sum takes over all the submatrices  $A_i$  constituted of  $i$  among the  $p + k$  columns of  $A$ . Hence, in order to minimize the expected number of eligible symbols  $E(A)$ , we choose  $A$  such that  $d_{\min}(A)$  is maximal, where  $d_{\min}(A)$  is the smallest number of linearly dependent columns of  $A$ .

1) *One extended bit per symbol node*: Consider the ensemble of regular  $(\lambda(x) = x, \rho(x) = x^3)$  LDPC codes defined over the  $\mathbb{F}_{16}$ . Assume that each symbol-node is extended by  $k = 1$  bit, such as to achieve an extended rate  $r_e = 0.4$ . According to the choice of the extended bit (among the 11 nontrivial extended bits),  $d_{\min}$  may be equal to 3, 4, or 5. The asymptotic threshold corresponding to each choice of the extended bit is shown in Figure 2. Note that extended bits are ordered on the abscissa according to the corresponding  $d_{\min}$ . For comparison purposes, we show also the asymptotic threshold corresponding to the repetition of some bit from the binary image (trivial extended bit  $\alpha_{2i} = x_i$ ), in which case  $d_{\min} = 2$ . Also, the blue line correspond to the threshold obtained if each symbol node was extended by choosing a random nontrivial extended bit. We observe that the best threshold is obtained when each symbol node is extended by  $\alpha_{15}$ , which is the XOR of the four bits  $x_0, x_1, x_2, x_3$  of the binary image.

2) *Several extended bits per symbol node*: We consider two ensembles of regular codes  $\mathcal{C}_1 (\lambda(x) = x, \rho(x) = x^3)$ ,  $\mathcal{C}_2 (\lambda(x) = x^2, \rho(x) = x^5)$  and one ensemble of semi-regular codes  $\mathcal{C}_3 (\lambda(x) = 0.5x + 0.5x^4, \rho(x) = 0.25x^4 + 0.75x^5)$ , of coding rate  $r = 1/2$ , defined over  $\mathbb{F}_{16}$ . For each ensemble of

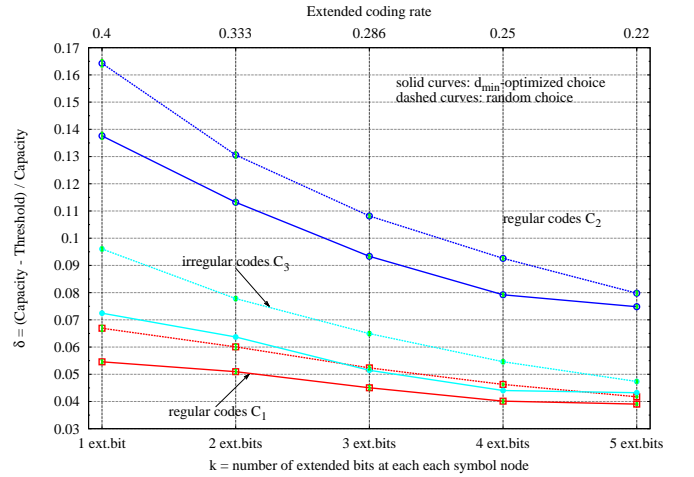


Figure 3.  $k$ -bits extension for regular and semi-regular NB-LDPC codes over  $\mathbb{F}_{16}$

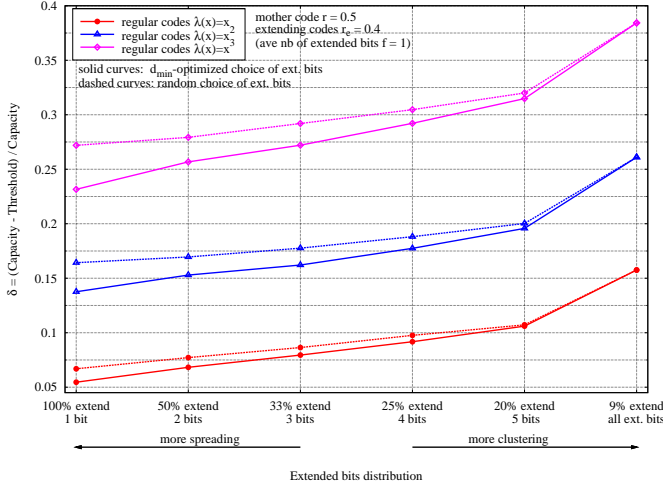


Figure 4. Comparison of spreading vs. clustering extending distributions for regular NB-LDPC codes over  $\mathbb{F}_{16}$

codes, we consider five different cases, in which all symbol nodes are extended by the same number  $k$  of bits, with  $k = 1, 2, 3, 4$ , and  $5$ . Accordingly, the extended coding rate  $r_e = 0.4, 0.33, 0.29, 0.25$ , and  $0.22$ .

The *normalized gap to capacity*, defined as:

$$\delta = \frac{\text{capacity} - \text{threshold}}{\text{capacity}} = \frac{1 - r - p_{\text{th}}}{1 - r},$$

is shown in Figure 3. Solid curves correspond to a  $d_{\min}$ -optimized choice of the extended bits, while dashed curves correspond to a random choice of the extended bits. For  $k = 5$ , there is only a small difference between these two strategies. However, when  $k \leq 4$ , the gain of the  $d_{\min}$ -optimized choice is significant for both regular and semi-regular codes.

### B. Extending distribution analysis

First of all, we discuss the case of regular codes. In Figure 4, we consider three ensembles of regular codes over  $\mathbb{F}_{16}$ , with coding rate  $r = 0.5$ . For each ensemble of codes, we consider five cases, corresponding to values of  $k$  between 1 and 5. In each case, a fraction  $f_k$  of symbol-nodes are extended by  $k$  bits, while the remaining symbol-nodes have no extended bit. The fraction  $f_k$  is chosen such that the extended coding rate  $r_e = 0.4$ . Hence,  $f_k = 1, 0.5, 0.33, 0.25, 0.20, 0.09$ , for  $k = 1, 2, 3, 4, 5$ , respectively. The right most point on the abscissa corresponds to a sixth case, in which the extended rate  $r_e = 0.4$  is achieved by extending 9% of symbol-nodes by  $k = 11$  bits (hence,  $k = q - p - 1$ , which is the maximum number of extended bits). For any of the three ensembles, we can observe that the smallest gap to capacity is obtained for  $k = 1$ , which means that extended bits are *spread* over as many symbol nodes as possible (in this case, 100%), instead of being *clustered* over the smallest possible number of symbol-nodes.

In case of irregular NB-LDPC codes, let  $\phi = \{f_{d,k}\}_{d,k}$  be an extending distribution. Thus,  $f_{d,k}$  is the fraction of degree- $d$  symbols with  $k$  nontrivial extended bits,  $0 \leq k < (q - p)$ . Let

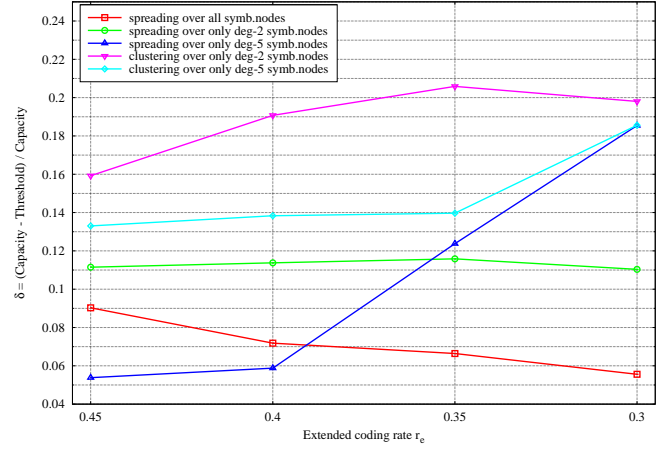


Figure 5. Comparison of several extending-distributions, for semi-regular NB-LDPC codes over  $\mathbb{F}_{16}$

$f_d$  denote the average number of extended bits per symbol-node of degree- $d$ ; that is:

$$f_d = \sum_{k=0}^{q-p-1} k f_{d,k} \in [0, q - p - 1]$$

We say that the *extending distribution*  $\phi$  is of *spreading-type* if for any degree  $d$ ,  $f_{d,k} \neq 0$  only if  $k = \lfloor f_d \rfloor$  or  $k = \lceil f_d \rceil$ . In different words, for any degree  $d$ , the extended bits are uniformly spread over all the symbol-nodes of degree  $d$ . Clearly, a spreading-type distribution is completely determined by the parameters  $\{f_d\}$ , as we have  $f_{d,\lfloor f_d \rfloor} = \lceil f_d \rceil - f_d$ ,  $f_{d,\lceil f_d \rceil} = \lfloor f_d \rfloor - f_d$ , and  $f_{d,k} = 0$  for  $k \notin \{\lfloor f_d \rfloor, \lceil f_d \rceil\}$ .

We say that the *extending distribution*  $\phi$  is of *clustering-type* if for any degree  $d$ ,  $f_{d,k} \neq 0$  only if  $k = q - p + 1$ . In different words, for any degree  $d$ , the extended bits are clustered over the smallest possible fraction of symbol-nodes of degree  $d$ . Clearly, a clustering-type distribution is completely determined by the parameters  $\{f_d\}$ , as we have  $f_{d,q-p+1} = \frac{f_d}{q-p+1}$  and  $f_{d,k} = 0$  for  $k \neq q - p + 1$ .

Now, let us consider the ensemble of semi-regular LDPC codes over  $\mathbb{F}_{16}$  with edge-perspective degree distribution polynomials  $\lambda(x) = 0.5x + 0.5x^4$  and  $\rho(x) = 0.25x^4 + 0.75x^5$ . The mother coding rate is  $r = 0.5$ , and we intend to extend symbol-nodes such as to achieve extended coding rates  $r_e \in \{0.45, 0.4, 0.35, 0.3\}$ . Several extending distributions are compared in Figure 5. There are three spreading-type distributions, which spread the extended bits over all the symbol-nodes, or only over the symbol-nodes of degree either 2 or 5, and two clustering-type distributions, which cluster the extended bits over the symbol-nodes of degree either 2 or 5. In all cases, extended bits (or, equivalently, extending matrices  $A_n$ ) are chosen such as to maximize the corresponding  $d_{\min}$  values. We observe that the smallest gap to capacity is obtained for extending distributions that spread extended bits either over the degree-5 symbol nodes only ( $r_e = 0.45, 0.4$ ), or over all the symbol-nodes ( $r_e = 0.35, 0.3$ ).



Table I  
OPTIMIZED EXTENDING DISTRIBUTIONS FOR A MOTHER NB-LDPC CODE  
WITH  $r = 0.5$  OVER  $\mathbb{F}_{16}$ , FOR  $r_e = \{0.45, 0.4, 0.35, 0.3, 0.25, 0.2\}$

$r_e$	0.5	0.45	0.4	0.35	0.3	0.25	0.2	2/15
$f_2$	0	0.4610	1.0164	1.7851	2.7442	4.1290	6.1737	11
$f_5$	0	0.3731	1.2113	1.2981	2.5055	3.5864	5.3409	11
$f_8$	0	0.2487	0.0359	1.8748	1.6831	2.3393	4.7494	11
$f_{18}$	0	0.1309	0.4871	0.8511	1.6415	2.9800	4.0234	11
$p_{th}$	0.4945	0.544	0.5939	0.6406	0.69	0.74	0.7872	0.8543
$\delta$	0.011	0.0109	0.0102	0.0145	0.0143	0.0133	0.016	0.0143
$\delta_{rand}$	0.011	0.0234	0.0284	0.0266	0.0251	0.0213	0.0172	0.0143

### C. Extending distribution optimization

Based on the above analysis, we only consider spreading-type extending distributions. Such an extending distribution is completely determined by the parameters  $\{f_d\}$ , and the extended coding rate can be computed by  $r_e = \frac{r}{1 + \frac{1}{p} \sum_d \Lambda_d f_d}$ , where  $\Lambda_d$  is the fraction of symbol-nodes of degree  $d$ .

For given degree distribution polynomials  $\lambda$  and  $\rho$ , and a given extending rate  $r_e$ , we use the differential evolution algorithm [11] to search for parameters  $\{f_d\}$  that minimize the asymptotic gap to capacity. We assume that, for each symbol-node, the extended bits are chosen such as to maximize the corresponding  $d_{min}$ . The optimized extended codes are presented in the next section.

### V. CODE DESIGN AND PERFORMANCE

In this section we present optimized extending distributions for an irregular mother code over  $\mathbb{F}_{16}$ . The mother code has coding rate  $r = 1/2$ , and it has been optimized by density evolution. The asymptotic threshold is  $p_{th} = 0.4945$ , and the edge-perspective degree distribution polynomials are:

$$\begin{aligned}\lambda(x) &= 0.596x + 0.186x^4 + 0.071x^7 + 0.147x^{17} \\ \rho(x) &= 0.2836x^4 + 0.7164x^5\end{aligned}$$

We optimized extending distributions for extended rates  $r_e \in \{0.45, 0.40, 0.35, 0.30, 0.25, 0.20\}$ . Optimized distributions  $\{f_d\}$  are shown in Table I, together with the corresponding asymptotic threshold  $p_{th}$  and normalized gap to capacity  $\delta$ . For comparison purposes, we have also indicated the normalized gap to capacity  $\delta_{rand}$  corresponding to a random choice of extended bits. The last column corresponds to extended rate  $r_e = 2/15$ , obtained by extending each symbol-node by the maximum number of extended bits, i.e.  $q - p - 1 = 11$  bits. It can be observed that the optimized distributions allow to maintain an almost constant value of  $\delta \approx 0.01$ , for all extended rates  $0.45 \geq r_e \geq 2/15$ .

Finally, Figure 6 presents the Bit Erasure Rate (BER) performance of optimized extending distributions for finite code lengths. All the codes have binary dimension (number of source bits) equal to 5000 bits (1250  $\mathbb{F}_{16}$ -symbols). The mother code with rate  $1/2$  has been constructed by using the Progressive Edge Growth (PEG) algorithm [12], and symbol nodes have been extended according to the optimized distributions (extension matrices  $A_n$  being chosen such as to maximize  $d_{min}(A_n)$ ).

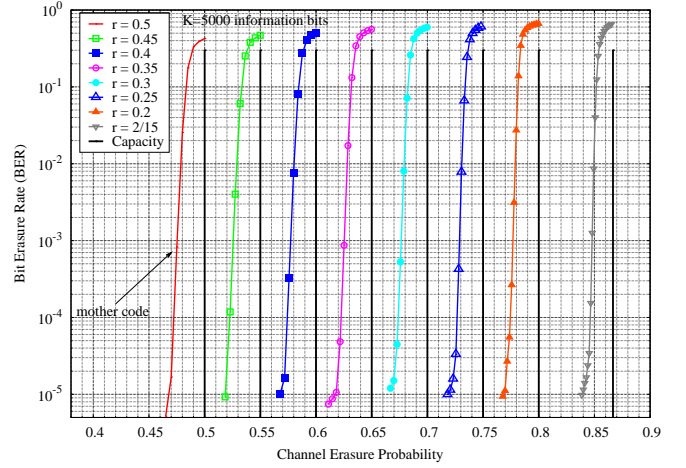


Figure 6. Finite length performance of optimized extended NB-LDPC codes

### VI. CONCLUSIONS

Based on the extended binary image of NB-LDPC codes, we presented a coding technique that allows to produce extra redundant bits, such as to decrease the coding rate of a mother code. The proposed method allows for using the same decoder as for the mother code: extra redundant bits transmitted over the channel are only used to “improve the quality of the decoder input”.

Extending distributions for regular and irregular codes have been analyzed by using simulated density evolution thresholds of extended codes over the BEC. We have also presented optimized extending distributions, which exhibit a normalized gap to capacity  $\delta \approx 0.01$ , for extended rates from 0.45 to  $2/15$ .

Finally, although this paper dealt only with NB-LDPC codes over the BEC, the results presented here can be easily generalized to different channel models.

### REFERENCES

- [1] M. Luby, “LT codes,” *Proc. ACM Symp. Found. Comp. Sci.*, 2002.
- [2] A. Shokrollahi, “Raptor codes,” *IEEE Transactions on Information Theory*, vol. 52-6, pp. 2551–2567, 2006.
- [3] R. G. Gallager, *Low Density Parity Check Codes*, M.I.T. Press, 1963.
- [4] V. Savin, “Binary linear-time erasure decoding for non-binary LDPC codes,” *IEEE Information Theory Workshop*, pp. 258–262, 2010.
- [5] K. Kasai and K. Sakaniwa, “Fountain coding with multiplicatively repeated non-binary LDPC codes,” in *IEEE Int. Symp. on Turbo Codes & Related Topics*, September 2010.
- [6] M. C. Davey and D. J. C. MacKay, “Low density parity check codes over GF(q),” *IEEE Commun. Lett.*, vol. 2, pp. 165–167, 1998.
- [7] R. M. Tanner, “A recursive approach to low complexity codes,” *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533–547, 1981.
- [8] T. J. Richardson and R. L. Urbanke, “The capacity of low-density parity-check codes under message-passing decoding,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599–618, 2001.
- [9] M. Gorgoglione, V. Savin, and D. Declercq, “Optimized puncturing distributions for irregular non-binary LDPC codes,” in *IEEE Int. Symp. on Information Theory and Applications*, 2010, pp. 400–405.
- [10] V. Savin, “Non-binary LDPC codes over binary erasure channel: density evolution analysis,” *IEEE Int. Symp. on App. Sci. on Biomed. and Comm. Tech.*, pp. 1–5, 2008.
- [11] R. Storn and K. Price, “Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces,” *J. of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [12] X. Y. Hu, E. Eleftheriou, and D. M. Arnold, “Regular and irregular progressive edge-growth tanner graphs,” *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 386–398, 2005.